Fast Algorithms for the Schur-Type Nonlinear Parametrization of Higher-Order Stochastic Processes

Agnieszka Wielgus, Jan Zarzycki
Signal Processing Systems Department
Wroclaw University of Science and Technology
Wroclaw, Poland
agnieszka.wielgus@pwr.edu.pl, jan.zarzycki@pwr.edu.pl

***Abstract*—We propose a class of fast algorithms, efficiently performing nonlinear Schur parametrization of higher-order and non-Gaussian stochastic processes, following from consideration of (weak) higher-order stationarity of the underlying signals and resulting in essential nonlinear complexity reduction, allowing for their practical implementations.**

***Index Terms*—Nonlinear Schur parametrization, nonstationary and stationary higher-order stochastic processes**

I. INTRODUCTION

One of the most important problems in signal processing is the Schur parametrization of stochastic signals [7], mapping a signal statistics into a set of the Schur coefficients which are used to solve a variety of applications-oriented topics (innovations filtering, decorrelation, whitening, stochastic modeling, pattern recognition (including speech and speakers recognition), transmission with compression of information (LPC method [9]), maximum entropy spectral estimation, Nevanlinna-Pick interpolation, among others). If an observed signal is Gaussian, its second-order statistics are the 'sufficient statistics'. Hence, the linear Schur parametrization algorithms are employed. In the non Gaussian case, the signal higher-order statistics have to be considered, and the linear approach has to be replaced by the nonlinear Schur-type parametrization algorithms whose complexity is tremendously growing (in comparison with the linear case) when the order as well as nonlinearity degree are updated, yielding (from some order and degree on) practically useless algorithms. Therefore, nonlinear complexity reduction is of crucial importance. To do that, in this paper we present a class of fast algorithms performing efficiently the nonlinear Schur parametrization of fourth-order stochastic processes (corresponding to the second degree nonlinear approach), following from a consideration of a class of fourth-order stationary processes. This approach, generalizing the corresponding linear case and resulting in a considerable complexity reduction, can be straightforwardly extended to higher degrees of nonlinearity, corresponding to consideration of higher-order stochastic processes.

II. LINEAR SCHUR PARAMETRIZATION OF SECOND ORDER STOCHASTIC PROCESSES

Let $\mathbf{y}$ denote a second-order, zero-mean, discrete-time and real-valued stochastic process observed on a finite time-interval and represented by the set of linearly independent random variables $\{y_t, y_{t-1}, \ldots, y_{t-n}\}$ where t is a reference point. The linear Schur parametrization procedure is a transformation, mapping the second-order statistics (e.g. covariance matrix) of $\mathbf{y}$ into a set of the so-called Schur coefficients, as it is schematically shown in Fig. 1. This transformation

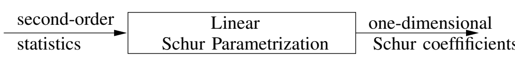

Fig. 1. The linear Schur parametrization problem.

is performed by the linear (J-orthogonal) innovations (equivalently: decorrelation or whitetning) filter whose parameters are exactly the extracted Schur coefficients, and - hence - the term 'parametrization'. This transformation results from the celebrated Schur algorithm [7] [5], and is closely related to the Levinson algorithm [4], yielding the theory of linear orthogonal filters [2].

A. Nonstationary case

1) Second-order nonstationary stochastic processes: Assume $t = 0$ and introduce the estimation space $S_0^n \triangleq \vee\{y_0, \ldots, y_{-n}\}$ (where $\vee\{\cdot\}$ stands for the 'span of'). Introducing the inner-product on S_0^n as $h_{i,k} = (y_{-i}, y_{-k}) \triangleq \mathbf{E}y_{-i}y_{-k}$ (where $\mathbf{E}$ is the expectation operator) the Gram (i.e. covariance) Hermitian positive-definite matrix of $\mathbf{y}$ will then be $H = [h_{i,k}]_{i,k=0,\ldots,n}$.

2) Linear Schur-type parametrization algorithms: The classical Schur parametrization is associated with second-order stationary processes. If a process is non-stationary, generalized Schur-type transformations have been introduced [6]. Here, we briefly recall the linear generalized Schur-type parametrization algorithm is a way suitable for nonlinear extensions. Consider a family of subspaces $S_i^k \triangleq \vee\{y_{-i}, \ldots, y_{-k}\}$ and define the forward estimate $\hat{y}_{i+1}^k \triangleq P(S_{i+1}^k)y_{-i} \in S_{i+1}^k$ (where $P(\cdot)$ is

ISBN 978-90-827970-1-5

the the orthogonal projection operator on $\vee\{\cdot\}$) together with the coprojection $\varepsilon_i^k \triangleq P(S_i^k \ominus S_{i+1}^k)y_{-i} \perp S_{i+1}^k$ (where $P(S \ominus S')$ stands for the orthogonal projection operator on the orthogonal complement of S' w.r.t. to S), and its nomalized version $e_i^k = \varepsilon_i^k \|\varepsilon_i^k\|^{-1}$. On the other hand, introduce the backward estimate $\breve{y}_i^{k-1} \triangleq P(S_i^{k-1})y_{-k} \in S_i^{k-1}$ together with the coprojection $\nu_i^k \triangleq P(S_i^k \ominus S_i^{k-1})y_{-k} \perp S_i^{k-1})y_{-k}$, and its normalized version $r_i^k = \nu_i^k \|\nu_i^k\|^{-1}$. Since $e_i^{k-1} \in S_i^{k-1}$ but $\perp S_{i+1}^{k-1}$ while $r_{i+1}^k \in S_{i+1}^k$ but $\perp S_{i+1}^{k-1}$, we have the following recurrence relations

$$\begin{bmatrix} e_i^k \\ r_i^k \end{bmatrix} = \theta(\rho_i^k) \begin{bmatrix} e_i^{k-1} \\ r_{i+1}^k \end{bmatrix} \tag{1}$$

where $\theta(\rho_i^k)$ is a J-orthogonal matrix:

$$\theta(\rho_i^k) \triangleq (1-(\rho_i^k)^2)^{-\frac{1}{2}} \begin{bmatrix} 1 & \rho_i^k \\ \rho_i^k & 1 \end{bmatrix} = \begin{bmatrix} ch\varphi_i^k & sh\varphi_i^k \\ sh\varphi_i^k & ch\varphi_i^k \end{bmatrix} \tag{2}$$

(with $\rho_i^k \triangleq (e_i^{k-1}, r_{i+1}^k)$ being the Schur coefficient of $\mathbf{y}$), and can be interpreted as a hyperbolic rotation matrix (an elementary parametrization section) if we put $\rho_i^k = th\varphi_i^k$, schematically shown in Fig. 2.

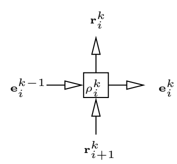

Fig. 2. Elementary linear hyperbolic rotation.

Connected accordingly together, those sections constitute the structure of the generalized (nonstationary) Schur parametrization scheme, being a cluster of nested hyperbolic rotations (shown in Fig. 3 for $n = 2$).

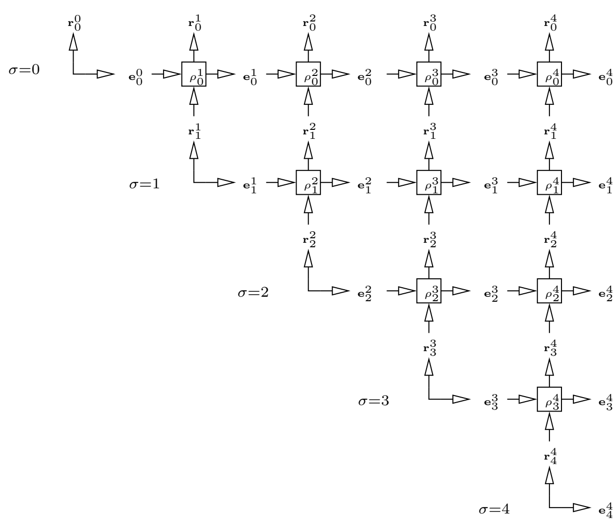

Fig. 3. Linear Schur parametrization of a 2-nd order nonstationary stochastic process ($n = 4$).

B. Stationary (in a weak second-order sense) case

1) Second-order stationary stochastic processes: Stationarity of the observed process results in the linear case in shift-invariance of the inner-product, i.e. $h_{i+\sigma,k+\sigma} = (y_{-i+\sigma}, y_{k+\sigma}) = \mathbf{E}y_{-i+\sigma}y_{k+\sigma} = \mathbf{E}y_{-i}y_k = (y_{-i}, y_k) = h_{i,k}$ (any σ) so that the second-order (two-dimensional) Hermitian covariance matrix H of the process becomes a Toeplitz matrix C.

2) Linear (classical) Schur parametrization algorithm: Hence, $\rho_{i+\sigma}^{k+\sigma} = -(e_{i+\sigma}^{k-1+\sigma}, r_{i+1+\sigma}^{k+\sigma}) = -(e_i^{k-1}, r_{i+1}^k) = \rho_i^k$, and we can restrict the algorithm only to the $\sigma = 0$ 'level' of the linear Schur parametrization procedure, as in the stationary case there is no nesting between the σ-levels anymore. This is illustrated in Fig. 4.

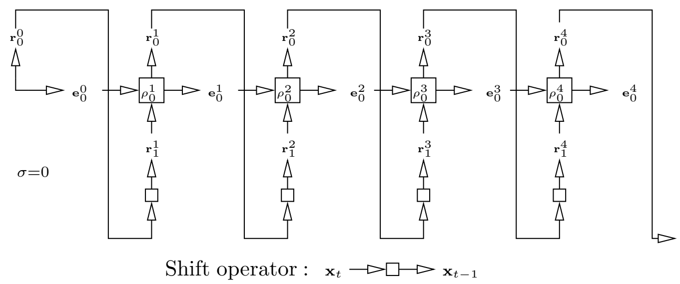

Fig. 4. Linear Schur parametrization of a 2-nd order stationary stochastic process ($n = 4$).

C. Linear complexity reduction

Second-order (weak) stationarity assumption results in major complexity reduction of the linear Schur paramerization algorithm in comparison with the general (Hermitian) case. The number of hyperbolic rotations in the nonstationary second-degree n-th order inear Schur parametrization scheme equals $N_{L_n} = \frac{2+n-1}{2}n$ while the associated hyperbolic rotations number in the stationary case is $S_{L_n} = n$. This linear complexity reduction is presented in Table 1.

Table 1: Linear complexity reduction

Order	No. of hyperbolic rotations		Compl. reduction
n	N_{L_n}	S_{L_n}	%
1	1	1	0
5	15	5	67
10	55	10	82
20	210	20	90
50	1 275	50	96
90	4 095	90	98

III. Nonlinear Schur parametrization of higher-order stochastic processes

To an $2M$-th order stochastic process (where for $M = 1$ we obtain the above presented linear parametrization algorithm for a second-order process) there corresponds an M-th degree nonlinear Schur-type parametrization procedure in which 'linear' and 'nonlinear' (multi-dimensional) Schur parameters are extracted (see Fig. III).

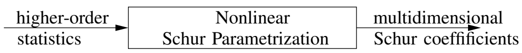

Fig. 5. The nonlinear Schur parametrization problem.

A. Nonstationary case

1) Higher-order nonstationary stochastic processes: Let us consider the following generalized (block, multi-indexed) covariance matrix of a fourth-order process

$$^{\{2\times 2\}}H = \begin{bmatrix} ^{1\oplus 1}H & ^{1\oplus 2}H \\ ^{2\oplus 1}H & ^{2\oplus 2}H \end{bmatrix} \tag{3}$$

where $^{1\oplus 1}H = [h_{ik}] = [(y_{-i}, y_{-k})]$, $^{1\oplus 2}H = [h_{ikl}] = [(y_{-i}, y_{-k}y_{-l})]$, $^{2\oplus 1}H = [h_{ijk}] = [(y_{-i}y_{-j}, y_{-k})]$, $^{2\oplus 2}H = [h_{ijkl}] = [(y_{-i}y_{-j}, y_{-k}y_{-l})]$. Actually, the matrix $^{\{2\times 2\}}H$ is the generalized Gram matrix of a fourth-order process (while $^{\{M\times M\}}H = [^{m\oplus u}H]_{m,u=1,\ldots,M}$ will be the generalized Gram matrix of a $2M$-th order process).

2) Nonlinear Schur-type parametrization algorithms: The Schur-type nonlinear parametrization of higher-order stochastic processes is actually equivalent to the generalized Gram-Schmidt orthogonalization of multi-indexed bases of three isomorphically isometric spaces [17], mapping the matrix H (3) into the generalized (block, multi-indexed) unit matrix, and yielding the set of the extracted generalized Schur coefficients. This work is inspired up to some extend by [16], although with a new – purely geometric – approach. Consider a simplest higher- (i.e. fourth-) order ($2M = 4$) process, associated with (non-trivial) second-degree nonlinear ($M = 2$) Schur parametrization problem and the underlying estimation space $S = \vee\{y_{-i}, y_{-i}y_{-j} \ ; \ _{i=0,\ldots,n\,;\,j=i,\ldots,n}\}$. Let $S_i \triangleq \vee\{y_{-i}, y_{-i}y_{-i}, \ldots, y_{-i}y_{-n}\}$. Then $S = S_0 \dot{+} \ldots \dot{+} S_n$ where $\dot{+}$ stands for direct sum of subspaces. Notice that in the nonlinear Schur-type parametrization (a generalized Gram-Schmidt orthogonalization) the following four types 'partial' forward and backward order-update recursions are considered: Linear-Linear (LL), Linear-Nonlinear (LN), Nonlinear-Linear (NL) and Nonlinear-Nonlinear (NN). Those recursions can be introduced in a similar way as in (1)-(2). To obtain - for example - the NN recursion, consider the subspace $S_{i,j}^{k,l} = \vee\{y_{-i}y_{-j}, \ldots, y_{-k}y_{-l}\}$ and define the forward estimate $\hat{y}_{i,j}^{k,l} \triangleq P(S_{i,j+1}^{k,l})y_{-i}y_{-j} \in S_{i,j+1}^{k,l}$ and the coprojection $\varepsilon_{i,j}^{k,l} \triangleq P(S_{i,j}^{k,l} \ominus S_{i,j+1}^{k,l})y_{-i}y_{-j} \perp S_{i,j+1}^{k,l}$ with its normalized version $e_{i,j}^{k,l} = \varepsilon_{i,j}^{k,l}\|\varepsilon_{i,j}^{k,l}\|^{-1}$. For the backward quatities, we obtain $\check{y}_{i,j}^{k,l} \triangleq P(S_{i,j}^{k,l-1})y_{-k}y_{-l} \in S_{i,j}^{k,l-1}$; $\nu_{i,j}^{k,l} \triangleq P(S_{i,j}^{k,l} \ominus S_{i,j}^{k,l-1})y_{-k}y_{-l} \perp S_{i,j}^{k,l-1}$ with its normalized version $r_{i,j}^{k,l} = \nu_{i,j}^{k,l}\|\nu_{i,j}^{k,l}\|^{-1}$, respectively. Since $e_{i,j}^{k,l} \in S_{i,j}^{k,l}$ but $\perp S_{i,j+1}^{k,l}$ while $r_{i,j}^{k,l} \in S_{i,j}^{k,l}$ but $\perp S_{i,j}^{k,l-1}$, we have the following recurrence relations

$$\begin{bmatrix} e_{i,j}^{k,l} \\ r_{i,j}^{k,l} \end{bmatrix} = \theta(\rho_{i,j}^{k,l}) \begin{bmatrix} e_{i,j}^{k,l-1} \\ r_{i,j+1}^{k,l} \end{bmatrix} \tag{4}$$

where
$\theta(\rho_{i,j}^{k,l})$ is the J-orthogonal matrix

$$\theta(\rho_{i,j}^{k,l}) = (1-(\rho_{i,j}^{k,l})^2)^{-\frac{1}{2}} \begin{bmatrix} 1 & \rho_{i,j}^{k,l} \\ \rho_{i,j}^{k,l} & 1 \end{bmatrix} = \begin{bmatrix} ch\varphi_{i,j}^{k,l} & sh\varphi_{i,j}^{k,l} \\ sh\varphi_{i,j}^{k,l} & ch\varphi_{i,j}^{k,l} \end{bmatrix} \tag{5}$$

(with $\rho_{i,j}^{k,l} \triangleq (e_{i,j}^{k,l-1}, r_{i,j+1}^{k,l})$ being the Schur coefficient of a fourth-order process $\mathbf{y}$), and can be interpreted as a hyperbolic rotation matrix (an elementary parametrization section) if we put $\rho_{i,j}^{k,l} = th\varphi_{i,j}^{k,l}$, schematically shown in Fig. 6 as the NN recursion d). The remaining hyperbolic rotations of Fig. 6 can be introduced in a similar way, i.e.:

a) the LL elementary section

$$\begin{bmatrix} e_i^k \\ r_i^k \end{bmatrix} = \theta(\rho_i^k) \begin{bmatrix} e_i^{l-1,l-1} \\ r_{i,j+1}^k \end{bmatrix} \tag{6}$$

$$\theta(\rho_i^k) = (1-(\rho_i^k)^2)^{-\frac{1}{2}} \begin{bmatrix} 1 & \rho_i^k \\ \rho_i^k & 1 \end{bmatrix} = \begin{bmatrix} ch\varphi_i^k & sh\varphi_i^k \\ sh\varphi_i^k & ch\varphi_i^k \end{bmatrix} \tag{7}$$

b) the NL elementary section

$$\begin{bmatrix} e_{i,j}^k \\ r_{i,j}^k \end{bmatrix} = \theta(\rho_{i,j}^k) \begin{bmatrix} e_{i,j}^{l-1,l-1} \\ r_{i,j+1}^k \end{bmatrix} \tag{8}$$

$$\theta(\rho_{i,j}^k) = (1-(\rho_{i,j}^k)^2)^{-\frac{1}{2}} \begin{bmatrix} 1 & \rho_{i,j}^k \\ \rho_{i,j}^k & 1 \end{bmatrix} = \begin{bmatrix} ch\varphi_{i,j}^k & sh\varphi_{i,j}^k \\ sh\varphi_{i,j}^k & ch\varphi_{i,j}^k \end{bmatrix} \tag{9}$$

c) the LN elementary section

$$\begin{bmatrix} e_i^{k,l} \\ r_i^{k,l} \end{bmatrix} = \theta(\rho_i^{k,l}) \begin{bmatrix} e_i^{k,l-1} \\ r_{i,j+1}^{k,l} \end{bmatrix} \tag{10}$$

$$\theta(\rho_i^{k,l}) = (1-(\rho_i^{k,l})^2)^{-\frac{1}{2}} \begin{bmatrix} 1 & \rho_i^{k,l} \\ \rho_i^{k,l} & 1 \end{bmatrix} = \begin{bmatrix} ch\varphi_i^{k,l} & sh\varphi_i^{k,l} \\ sh\varphi_i^{k,l} & ch\varphi_i^{k,l} \end{bmatrix} \tag{11}$$

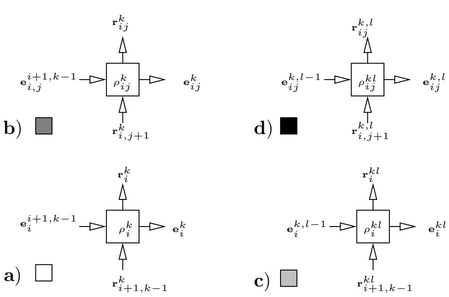

Fig. 6. Elementary linear/nonlinear hyperbolic rotations: a) LL (Linear-Linear), b) NL (Nonlinear-Linear), c) LN (Linear-Nonlinear), d) NN (Nonlinear-Nonlinear).

Connected accordingly together, they constitute a block nonlinear Schur parametrization (generalized Gram-Schmidt orthogonalization) section schematically shown in Fig. 7 where at each block-step a pair of 'new' partial forward and backward errors is updating the scheme.

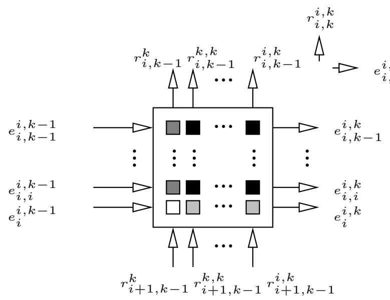

Fig. 7. The block-section of elementary linear/nonlinear hyperbolic rotations.

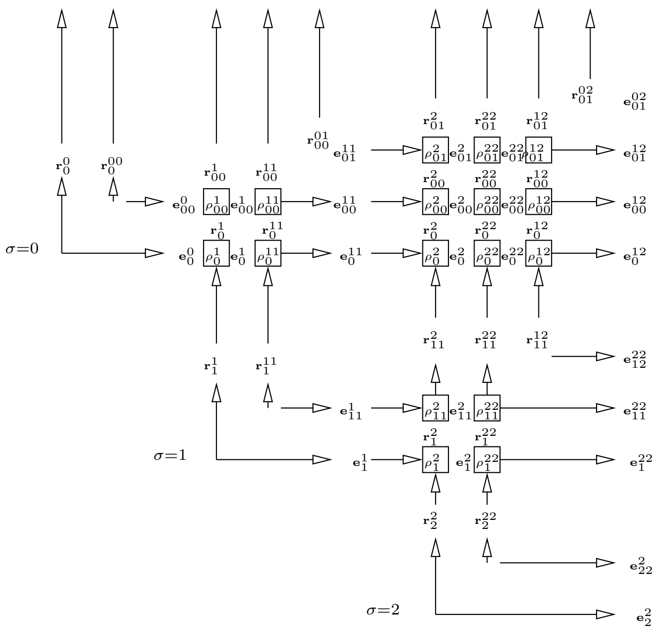

Fig. 8. The detailed structure of the nonlinear Schur parametrization of a 4-th order nonstationary stochastic process ($n = 2$).

B. Stationary (in a weak higher-order sense) case

1) Higher-order stationary stochastic processes: The consequence of the (weak) fourth-order stationarity of the observed process is shift-invariance of the inner-products (for any σ): $h_{i+\sigma,k+\sigma} = (y_{-i+\sigma}, y_{-k+\sigma}) = \mathbf{E}y_{-i+\sigma}y_{-k+\sigma} = \mathbf{E}y_{-i}y_{-k} = (y_{-i}, y_{-k}) = h_{ik}$; $h_{i+\sigma,k+\sigma,l+\sigma} = (y_{-i+\sigma}, y_{-k+\sigma}y_{-l+\sigma}) = \mathbf{E}y_{-i+\sigma}y_{-k+\sigma}y_{-l+\sigma} = \mathbf{E}y_{-i}y_{-k}y_{-l} = (y_{-i}, y_{-k}y_{-l}) = h_{ikl}$; $h_{i+\sigma,j+\sigma,k+\sigma} = (y_{-i+\sigma}y_{-j+\sigma}, y_{-k+\sigma}) = \mathbf{E}y_{-i+\sigma}y_{-j+\sigma}y_{-k+\sigma} = \mathbf{E}y_{-i}y_{-j}y_{-k} = (y_{-i}y_{-j}, y_{-k}) = h_{ijk}$; $h_{i+\sigma,j+\sigma,k+\sigma,l+\sigma} = (y_{-i+\sigma}y_{-j+\sigma}, y_{-k+\sigma}y_{-l+\sigma}) = \mathbf{E}y_{-i+\sigma}y_{-j+\sigma}y_{-k+\sigma}y_{-l+\sigma} = \mathbf{E}y_{-i}y_{-j}y_{-k}y_{-l} = (y_{-i}y_{-j}, y_{-k}y_{-l}) = h_{ijkl}$; so that the generalized Hermitian covariance matrix ${}^{\{2\times2\}}H$ (3) becomes a generalized (block, multi-indexed) Toeplitz matrix ${}^{\{2\times2\}}C$.

2) Nonlinear Schur-type parametrization algorithm: From the above properties of higher-order processes it clearly follows that in the nonlinear stationary case we obtain $\rho_{i+\sigma}^{k+\sigma} = \rho_i^k$, $\rho_{i+\sigma}^{k+\sigma,l+\sigma} = \rho_i^{kl}$, $\rho_{i+\sigma,j+\sigma}^{k+\sigma} = \rho_{ij}^k$ and $\rho_{i+\sigma,j+\sigma}^{k+\sigma,l+\sigma} = \rho_{ij}^{kl}$ regardless of the σ-shift. Taking in mind that the initializations (i.e. the partial backward prediction errors, constituting the ON basis of the estimation space) are actually the shifted versions of the outcomes of the 'upper-wire' of the previous block-step (due to the shift-invariance of the inner-product), we can confine the algorithm to the $\sigma = 0$ 'level' of the nonlinear Schur parametrization scheme only, as in the higher-order stationary case there is no nesting between the σ-levels anymore (similarly as in the linear stationary case). This is illustrated in Figs. 9 and 11. Let us mention that the complexity reduction implied by the higher-order stationarity can directly be generalized to $2M$-th order processes, and M-th degree nonlinear Schur-type parametrization algorithms.

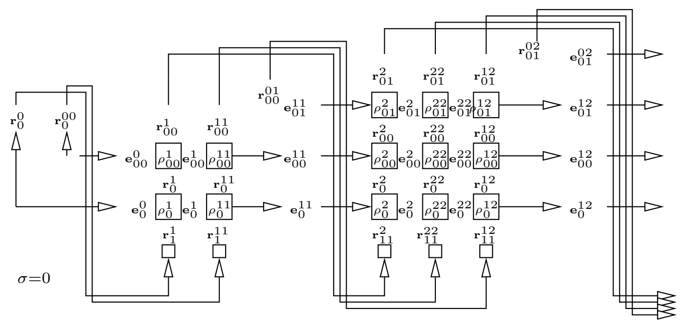

Fig. 9. The detailed structure of the nonlinear Schur parametrization of a 4-th order stationary stochastic process ($n = 2$).

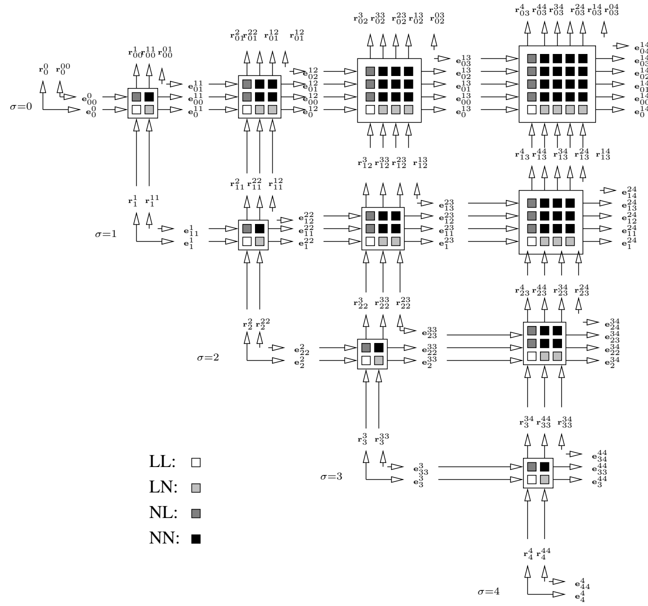

Fig. 10. The block structure of the nonlinear Schur parametrization of a 4-th order nonstationary stochastic process ($n = 4$).

C. Nonlinear complexity reduction

The number of hyperbolic rotations in the stationary second-degree n-th order nonlinear Schur parametrization scheme equals $S_{N_n} = \frac{1}{6}n(n+1)(2n+1)$ while the associated hyperbolic rotations number in the nonstationary case is

ISBN 978-90-827970-1-5

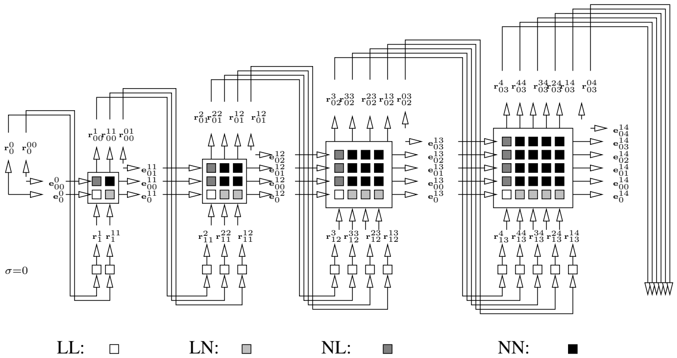

Fig. 11. The block structure of the nonlinear Schur parametrization of a 4-th order stationary stochastic process ($n = 4$).

$N_{N_n} = \sum_{i=1}^{n} S_n$. The nonlinear complexity reduction is presented in Table 2.

Table 2: Nonlinear complexity reduction

Order	No. of hyperbolic rotations		Complexity reduction
n	N_{N_n}	S_{N_n}	%
1	1	1	0
5	105	55	48
10	1 210	385	68
20	16 170	2 870	82
50	235 340	22 140	92
90	5 713 890	241 065	96

IV. Conclusion

The above results compare the Schur-type parametrization algorithms for the nonstationary case (when higher-order covariances are actually generalized Hermitian matrices) versus the stationary situation (when those matrices become generalized Toeplitz). In real life, however, we are often faced with near-stationary signals (a good example is speech signal) whose covariance matrices are 'close' to Toeplitz (in a well-defined sense) for which Schur parametrization algoritms complexity is higher than in the stationary case, but considerably reduced comparing to a 'totally' nonstationary case. Two approaches for second-order processes have been introduced and are mentioned here: the approach based on the notion of α-stationarity (see e.g. [8]) and the concept based on the staircase matrix extension problem (see e.g. [3]). The second approach has been generalized to higher-order processes in [17], [18], [19] while the first one was our inspiration for introducing of a class of 'p-stationary' second-order processes [21], [22] and the associated Schur parametrization and modeling algorithms with reduced complexity. The 'p-stationary' approach can straightforwardly be generalized for higher-order processes with statistics being 'close' to generalized Toeplitz matrices. Taking an advantage of higher-order closeness to the Toeplitz case, the corresponding class of nonlinear algorithms (attractive in implementations) is the present direction of our current research.

References

[1] P.Dewilde, E.F.A.Deprettere, *The Generalized Schur Algorithm: Approximation and Hierarchy*, in: **Operator Theory: Advances and Applications**, vol. 29, Birkhäuser Verlag, Basel, 1988, pp.97-116.

[2] P.Dewilde, *Stochastic Modelling with Orthogonal Filters*, in: **Outils et modeles mathematiques pour l'automatique, l'analyse de systemes et le traitement du signal**, CNRS (ed.), Paris 1982, pp.331-398.

[3] P.Dewilde, *A Course on the Algebraic Schur and Nevanlinna-Pick Interpolation Problems*, in: **Algorithms and Parallel VLSI Architectures**, vol.A: Tutorials, E.F.Deprettere and A.-J. van der Veen (eds.), Elseviere Science Publ., 1991, pp.13-69.

[4] P.Dewilde, A.C.Vieira, T.Kailath, *On a Generalized Szegö-Levinson Realization Algorithm for Optimal Linear Predictors Based on a Network Synthesis Approach*, IEEE Trans. on Circuits and Systems, vol. CAS-25, No.9, September 1978, pp.663-675.

[5] P.Dewilde, H.Dym, *Schur Recursions, Error Formulas and Convergence of Rational Estimators for Stationary Stochastic Sequences*, IEEE Trans. on Information Theory, vol. IT-27(4), July 1981, pp.446-461.

[6] P.Dewilde, E.F.A.Deprettere, *The Generalized Schur Algorithm: Approximation and Hierarchy*, in: **Operator Theory: Advances and Applications**, vol. 29, Birkhäuser Verlag, Basel, 1988, pp.97-116.

[7] T.Kailath, *A Theorem of I.Schur and Its Impact on Modern Signal Processing*, in: **I.Schur Methods in Operator Theory and Signal Processing**, I.Gohberg (Ed.), Operator Theory: Advances and Applications, vol.18, Birkhäuser-Verlag 1986, pp.9-30.

[8] T.Kailath, *Linear Estimations for Stationary and Near-Stationary Processes*, in: **Modern Signal Processing**, T.Kailath (Ed.), Hemisphere Publishing Corp./Springer-Verlag 1985, pp.59-128.

[9] P.Kroon, E.F.A.Deprettere, R.J.Sluyter, *Multi-pulse excitation linear-predictive speech coder*, US Patent 4932061, 1990.

[10] H.Lev-Ari, T.Kailath, *Lattice filter parametrization and modelling of nonstationary processes*, IEEE Trans. on Inf. Theory, vol. IT-30, 1984, pp.2-16.

[11] M.Schetzen, *Volterra-Wiener theories of nonlinear systems*, Wiley, New York, 1980.

[12] N.Wiener, *Nonlinear problems in random theory*, Mass. Inst. Techn. Press, New York: Wiley, 1958, Technology Press.

[13] U.Libal, A.Wielgus, W.Magiera, *Nonlinear Orthogonal Parametrization and Modeling Algorithms for Higher-Order Non-Gaussian Time-Series*, IEEE Signal Processing Symposium (SPSympo-2017), special session on Stochastic Realization and Orthogonal Signal Processing.

[14] J.Zarzycki, P.Dewilde, *The nonlinear nonstationary Schur algorithm*, Proc. Workshop on Advanced Algorithms and Their Realizations, Editor M.Verhaegen (Delft Univ. Techn., Chateaux de Bonas, 1991, paper V3.

[15] J.Zarzycki, A.Wielgus, U.Libal, *Nonlinear Schur-Type Orthogonal Transformations of Higher-Order Stochastic Processes: An Overview of Current Topics*, IEEE Signal Processing Symposium (SPSympo-2017), special session on Stochastic Realization and Orthogonal Signal Processing.

[16] J.Zarzycki, *Nonlinear least-squares prediction filters for higher-order stochastic sequences*, Lecture Notes in Control and Information Sciences, 1985, Vol. 73, Springer-Verlag, Boston/Heidelberg/NY/Tokyo.

[17] J.Zarzycki, *Multidimensional Nonlinear Schur Parametrization of Non-Gaussian Stochastic Signals – Part One: Statement of the Problem*, MDSSP Journ., July 2004, Volume 15, Issue 3, pp. 217-241.

[18] J.Zarzycki, *Multidimensional Nonlinear Schur Parametrization of Non-Gaussian Stochastic Signals – Part Two: Generalized Schur Algorithm*, MDSSP Journ., July 2004, Volume 15, Issue 3, pp. 243-275.

[19] J.Zarzycki, *Multidimensional Nonlinear Schur Parametrization of Non-Gaussian Stochastic Signals – Part Three: Low-Complexity Solution*, MDSSP Journ., October 2004, Volume 15, Issue 4, pp. 313-340.

[20] A.Wielgus, U.Libal, W.Magiera, *Nonlinear Complexity Reduction: Sparsity of the Generalized Schur Coefficient Matrices and Frobenius Norm Criterion*, Signal Processing Symposium (SPSympo-2017), special session on Stochastic Realization and Orthogonal Signal Processing.

[21] A.Wielgus, J.Zarzycki, *Efficient Schur Parametrization of Near-Stationary Stochastic Processes*, IEEE IWSSIP-2017.

[22] A.Wielgus, J.Zarzycki, F.Lwow, *Schur Parametrization and Orthogonal Modeling of p-Stationary Second-Order Stochastic Processes*, submitted for publication.